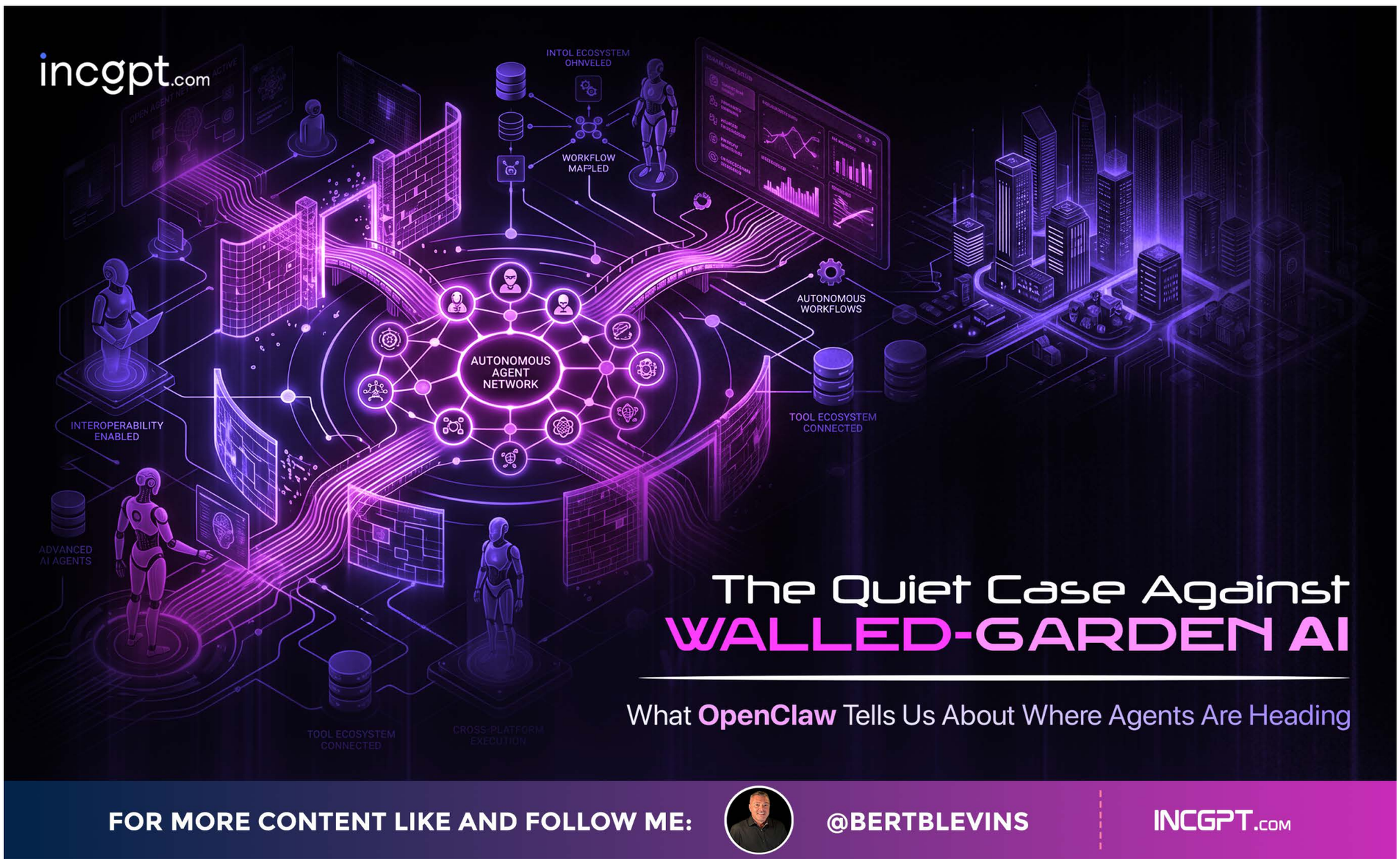
The Quiet Case Against Walled-Garden AI:

What OpenClaw Tells Us About Where Agents Are Heading

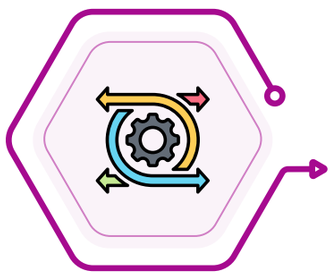


An open-source project, a wave of Chinese clones, and even an Nvidia release are quietly proving that AI agents do not have to belong to a single platform.

AI agents have arrived inside the products people already use every day. Google's Gemini can now reach into a phone's apps to place a food order. Microsoft's Copilot can spin up a full PowerPoint deck from a short prompt. On the surface, it looks like a familiar story: the same handful of dominant platforms extending their reach into yet another layer of digital life. But a separate project is sketching out a very different path. OpenClaw, an open-source AI agent, is showing that an agent does not have to be welded to a single ecosystem to be useful.

Meet OpenClaw: one agent, many brains

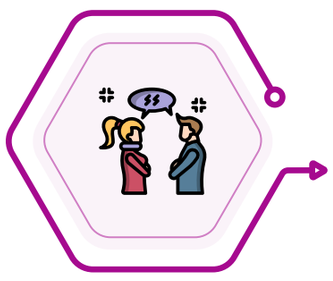
OpenClaw is the work of Austrian developer Peter Steinberger. Its defining feature is that it is not bound to any one foundation model. A user can plug in Anthropic's Claude, OpenAI's ChatGPT, or an open-weight model such as DeepSeek, and the change takes nothing more than a one-line command in the interface. The agent does not need to be reintroduced to its user each time.



That flexibility is possible because of how the system is built. At its core sits a piece of software called the Gateway, which runs on the user's own device. The Gateway is what actually connects to external services like email and calendar, and it also holds the user's memory and preferences, all stored locally rather than on a vendor's servers. When the user gives the agent a task, such as booking a restaurant, the Gateway hands the relevant context — calendar availability, dietary preferences, contacts — to whichever model is doing the reasoning. Because the personal layer lives on the device, switching the underlying model does not mean starting from scratch.

A wave of copies, including one from Nvidia

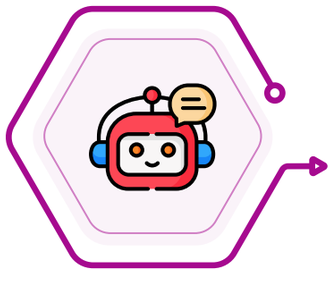
OpenClaw has not stayed a niche curiosity. It has triggered a rush of similar products, especially among China's largest technology firms. Xiaomi released Miclaw, Moonshot AI launched Kimi Claw, and Zhipu AI shipped AutoClaw. Nvidia has now joined the same camp with NemoClaw, an open-source release built directly on the OpenClaw framework, layered with extra security and privacy tooling. NemoClaw defaults to Nvidia's own Nemotron models but still lets users swap in another foundation model if they prefer.



Taken together, these projects make an argument by example. An AI agent is not a single indivisible product. The foundation model and the surrounding software layer are genuinely separable, which means a more modular agent market is technically possible. Users could enjoy meaningful choice at each layer rather than picking one platform and inheriting whatever model, integrations, and policies come bundled with it.

Why the big platforms are pulling in the opposite direction

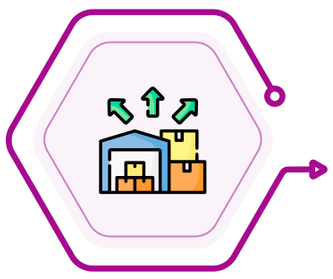
The leading commercial AI players are not heading toward this modular future. They are doing the opposite, steadily widening the scope of what their agents can do inside their own ecosystems. Gemini can now search through a user's photos and read their group chats. ChatGPT can shop on a user's behalf, and recently gained the ability to connect to health apps and medical records. Each capability looks genuinely useful in isolation, but it also pulls more of the user's life inside a single provider's perimeter.



That matters because chatbot conversations already contain unusually intimate information. Layered with calendars, messages, photos, purchases, and health data, an agent ends up holding the kind of profile that once required entire industries of data brokers and tracking cookies to assemble. Concentrating it inside one company makes targeted advertising more precise and more profitable, and OpenAI has already begun testing ads inside ChatGPT.

Bundling, walled gardens, and the temptation to self-preference

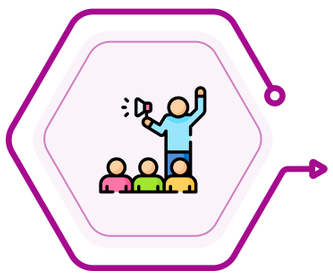
Distribution amplifies the effect. Google has spread Gemini across Android, Search, and the Workspace suite. Microsoft tucked Agent Mode into Copilot, which itself rides inside Office 365. For most users, these agents are not so much chosen as inherited along with the platforms they already pay for. That is a powerful position for any vendor to occupy.



It also opens the door to a familiar problem: self-preferencing. A company that builds, distributes, and operates an agent has obvious reasons to nudge users toward its own products and partners. Search engines have already lived through this debate; Google was fined €2.4 billion by European regulators for burying rival shopping services below its own in search results. An agent can do something similar, only less visibly. A traditional search results page at least shows several options. An agent often offers a single recommendation, with no clear window into how it picked one over the others.

How lock-in quietly compounds

The same personal data that powers helpful agents also becomes a quiet form of lock-in. Over time, an agent learns a user's routines, tastes, and shortcuts, including both what the user has stated outright and what the agent has inferred. Walking away from a vendor means rebuilding that knowledge somewhere else and reconnecting every app, file, and service. Every additional week of use raises the cost of leaving.



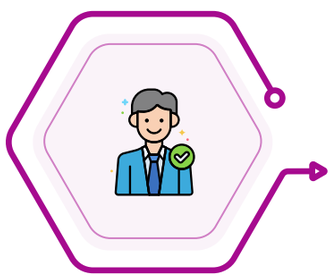
OpenClaw, and the projects following in its wake, are a working demonstration that this lock-in is a design choice rather than a law of physics. Activity logs and memories are generated automatically and stored on the user's device in human-readable form, meaning the user can edit them, delete them, or move them. Integrations such as email and calendar persist when the underlying model changes. The personal layer is not held hostage by any single provider.

A side benefit: less platform surveillance

Modularity has a privacy dividend as well. If a user can rotate between foundation model providers, no single company ends up with a complete picture of that person's online behavior. For genuinely sensitive tasks, users can fall back to smaller open-weight models that run entirely on their own hardware, ensuring that no data leaves the device at all. That is a structural form of privacy protection, not just a policy promise.

The security trade-off is real

Modularity is not free. Security researchers have flagged a number of risks tied to OpenClaw's access to sensitive personal data and its exposure to untrusted content. Attackers have already taken advantage. One report identified hundreds of malicious add-ons aimed at stealing user data, and prompt injection attacks, where a hidden instruction lurking in an email or webpage hijacks the agent's behavior, remain an open problem.



In a vertically integrated agent, a single provider can at least take central responsibility for securing the entire stack. In a modular setup, more of that responsibility shifts to the user. Closing the gap will require new tooling and safeguards, which are starting to emerge in the form of dedicated security startups and add-ons.

Why this will not happen without rules

There is one stubborn obstacle to the modular vision: today's platforms do not benefit from it. Meta showed as much when it blocked rival AI assistants from operating inside WhatsApp in favor of its own. The Italian Competition Authority and the European Commission stepped in quickly with interim measures. Meta's move was unusually visible, but a platform that simply never opens up in the first place achieves the same lock-out without inviting any regulatory attention at all. Unless gatekeepers are required to provide meaningful access to their platforms, alternatives like OpenClaw can demonstrate what is possible without ever reaching the mainstream.

A blueprint, even if its creator has moved on

There is some irony in the story. Steinberger himself has since been hired by OpenAI, one of the companies most associated with the vertically integrated approach his project pushed against. Yet OpenClaw remains a working template for agent portability, where integrations and memories survive a switch between foundation models. The argument it implies is that mainstream agents should be held to a similar standard. At a minimum, users should have a clear right to export their data — memories, chat histories, uploaded files — in standardized, human-readable formats so that moving providers does not mean starting their digital lives over.

The current shape of the agent market can feel like a foregone conclusion, with a few familiar names racing to lock users inside ever-larger walled gardens. OpenClaw and the projects it has inspired suggest otherwise. The market is still young, the technical separation between layers is real, and the design choices that determine whether users end up in control or simply enclosed have not all been made yet.